BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GENERAL SCANNING INC.
Laser Systems Division

Laser Link Fusing

Guidelines for the Applications of M3XX Laser Systems
in Semiconductor Link Fusing Processes

General Scanning Inc.
Laser Systems Division
32 Cobble Hill Road
Somerville, MA 02143

January 1997

Table of Contents

Introduction

The preferred means of implementing redundancy in large memories is with laser ablation of polysilicon or other link structures. Laser fusing of links on semiconductor devices has been in widespread commercial use since the late 1970's. All known commercial 4M and larger DRAM devices are implemented with laser fusing capability for repair. The yield of other memory devices such as SRAMs, some Flash devices, EPROM's and even mask ROMs are commonly enhanced using laser redundancy. Since the early 1990's processor-specific cache controller devices have achieved memory densities that benefit from the use of redundancy implemented with a laser.

The dominant commercial applications of lasers (of the class used for memory repair) is in the process of repairing memory devices. Similar laser technology is also used to program certain logic devices by removing links. The number of links removed per die can be very large, otherwise the process is similar to memory repair. A special sub-class of logic programming is the serializing of RF ID transponders by laser removal of links. Some sensors and other linear and mixed-signal devices are "tuned" by selective removal of metal or other links (this is distinct from widely used *thin-film* laser trimming process).

LSD has a *family* of products that are frequently used for laser applications in the semiconductor industry. The M3 product family has a common base of modules (laser, beam positioning technology, software, etc.) with distinct optical features and specifications for specific applications. The M320 and M310 member of the M3 product family are commonly used for laser removal of links. The M320 is preferred for virtually all memory repair applications due to the large field size, small minimum spot size and extreme accuracy of the beam positioner. The M310 is often used for link removal on sensors, RF ID transponders and some logic devices. The M310 has a smaller field size, larger minimum spot size, lower accuracy and lower cost when compared with the M320.

This document provides answers to frequently asked questions about the use of LSD laser systems in memory repair and related laser applications. The emphasis of the document is on how to apply an M3 laser system to a practical application, **not** on how the M3 systems operate. Suggestions are made for the design of devices to maximize device yield and improve the reliability of the laser process in a typical manufacturing environment.

Suggestions on how the document can be improved are welcome.

Proprietary Information

This document contains proprietary information that is the property of General Scanning Inc., Laser Systems Division. This document is provided to existing and prospective customers of LSD for the sole purpose of assisting users in the application of LSD laser system products in their manufacturing processes.

This document must be handled in accordance with stipulations of any Non-disclosure Agreement existing between LSD and the recipient covering proprietary information.

Draft Revision 1: This document is in the process of being updated with recent application experiences. Draft Revision 1 is provided on an "as is" basis while the update is in process. A formal revised copy will be provided to replace the Draft copy when it becomes available.

Link Structures for Laser Fusing

The specific composition of most link structures is regarded as proprietary by most manufacturers. LSD operates to maintain customer-provided information in confidence when such information is provided to LSD under a mutual non-disclosure agreement. Virtually all manufacturers work with LSD under such an agreement when developing new processes.

However, general information is widely available in the industry about "generic" link structures. The following suggestions are offered as a starting point for process development. The precise structures described are not known to exist on any particular device. They reflect commonly known structures and industry-wide practice as required by typical semiconductor processes, design rules and structures reported in the literature.

Most current generation memory devices implement fusible links with a silicide structure incorporating silicon sintered with one of the following metals: Tungsten (W), Titanium (Ti) or Molybdenum (Mo). Polysilicon was widely used in the past, but silicide links are preferred due to higher conductivity.

An ideal link structure is less then 10,000 Å thick consisting of silicide materials or a sandwich structure of metals and antireflective coatings. The insulator over the link may consist of differing types of oxide such as SOG and BPSG as well as silicon nitride. These oxides layers should not be greater than 10,000 Å thick. Sometimes the passivation oxide over the link may be etched back to reduce the oxide thickness. However providing the correct etch stop may be difficult or require a number of extra processing steps. If thick oxides are used then link removal may become difficult and multiple laser pulses may be required to achieve optimum yield.

Metal link structures are under active investigation as fusible links due to better conductivity and improved registration of the links to laser alignment fiducials - see discussion below. Also, with high density memories using multilayer interconnects, the metal layers are closer to the top of the structure. Metal links have the advantage of not requiring as much etching of the oxide over the links as silicide structures.

The laser fusing event is a complex thermodynamic process that is difficult to model analytically. The principal work that has been done to understand the process can be found in a paper by Chilipala and Scarfone ("Computer Simulation of Target Link Explosion in Laser Programmable Redundancy for Silicon Memory", Journal for Material Research, 1 (2) March/April 1986). A simplified explanation follows. At the onset of the laser pulse incident on a link structure, the link material absorbs a fraction of the photon energy based on the room temperature absorbtivity of the material at the laser wavelength (1.047μ). The energy is absorbed as heat which causes the temperature in the region local

to the incident radiation to rise. Most link structures exhibit a temperature dependent absorbtivity at the incident wavelength. As the line temperature rises, the fraction of the pulse that is absorbed increases. This created a "thermal runaway" condition since the increased absorbtivity causes more energy to be converted to heat which raises the temperature further which increases the absorbtivity, etc. As the link temperature rises, the pressure under the passivation local to the hot spot increases rapidly until the passivation ruptures. The (molten) link material is explosively ejected from the structure when the pressure is released by the rupture of the passivation.

The quality of the fusing process can be partially assessed by observing those factors that have been shown over time to correlate with device yield and long term reliability. Only extensive testing and long term stress testing, however, will uncover a subtle fusing problem A good quality cut will exhibit minimal residual link material in the vicinity of the "crater" . The oxide over the link will not show signs of cracking beyond a region local to the link and the crater will not extend into the adjoining metal runners or any other structures. There should not be any cracking of the oxide or damage to the substrate beneath the removed link. There should be no visible effect of the pulse on any adjacent links.

Much of the effort in laser process development is focused on achieving a high quality laser cut. Minor changes in device geometry, link and oxide composition and dimensions as well as laser parameters will affect the cut quality. LSD is frequently involved in the process development efforts with customers: developing test protocols, setting laser parameters, developing novel cutting methods and even modifying the laser energy distribution, temporal pulsewidth and other parameters.

Metal links have received much attention recently due to the attendant advantages: better conductivity, less oxide etching and the prospect of placing alignment fiducials in the same layer as the links to be removed. The first benefit is increasingly important as device speeds continue to increase and the layout of redundant cells require propagation paths through more links. Metal interconnect layers are covered by thinner oxide layers than silicide structures. Thinner oxide is preferred for laser fusing. Improved registration is important in achieving optimum accuracy of spot placement. Reducing error in spot placement allows for increased process latitude, smaller link pitch and improved yield.

Metal fuse structures are not widely used as of this writing except in special situations. Metal is generally more difficult to fuse, produces a less clean cut and may require additional steps after the laser fusing process to achieve the desired level of cut quality. Work on metal fuses done by LSD in conjunction with a local university has shown great promise. Metal structures were designed and fused with extremely good cut quality. The metal structures were treated with an anti-

reflective coating and the links were less then 7500 Å thick. The oxide over the links was also very thin (4000 Å).

Link geometry may have a profound impact on device yield. The ideal link from laser considerations is long, narrow and spaced on a wide pitch. The ideal link from the device designer standpoint is short, wide and spaced on a fine pitch. The short wide link has better conductivity than the long narrow design and the real-estate savings from a fine pitch are obvious. A more detailed analysis of the tradeoffs between pitch and real-estate is saved for the section on spot size below. Typical non-metal links are a maximum of 1.2μ wide, a minimum of 8μ long and spaced on 5μ or greater pitch.

Link Location Considerations

The placement of links can significantly affect throughput when a large number of links must be removed. System throughput is optimized when the Linksprint features of the M3XX system is used. Linksprint sweeps the laser spot across a group of links at a constant rate while firing the laser to remove selected links. The beam motion and laser pulsing are precisely controlled by a dedicated DSP processor in the M3 control electronics to insure accurate placement of laser pulses. The laser is fired at every link, however, the acousto-optic modulator (AOM) is kept closed to prevent the laser pulses from reaching the links. The AOM is opened only for selected laser pulses to allow the laser energy to pass through the beam positioned and remove a link.

The physics of the laser requires that the laser be operated at a fixed frequency for frequencies above approximately 500 Hz in order to have consistent laser energy per pulse. Linksprint operates at frequencies up to almost 4000 Hz; therefore the laser is pulsed at a fixed rate. The inertia of the beam steering components prevents the beam positioner from changing speed instantaneously. The beam speed, laser firing rate and link pitch are interrelated.

Beam Speed = Laser Rate x Link Pitch

Since beam speed and laser rate must be kept constant due to the laws of physics, then the link pitch must also be kept constant.

The ideal die, from a laser throughput standpoint, would have all the links on the die arranged in a single long row all on a common pitch. Practical device layouts incorporate links on a common pitch in small groups; but distribute the groups at random throughput the die due to non-laser related considerations. The requirement for common link pitch holds true only within a single group. Distinct links groups may have distinct link pitches. Distinct link groups can be arranged in any direction (X or Y).

The choice of link pitch affects the cutting speed. The maximum beam speed is limited to approximately 50,000μ / sec. The maximum laser pulse rate is limited to approximately 4000 pulses per second. The actual link processing rate is constrained by one of these two limits. At pitches less than 50,000 / 4000 = 12.5μ, the speed is laser rate limited. Therefore the cutting speed is about 4000 links per second. Any reduction in link pitch below 12.5μ will have no effect on throughput.

At link pitches greater than 12.5μ, the maximum beam speed determines the actual link cutting rate. For instance, links spaced on a 20μ pitch will be cut at a rate of (approximately) 50,000 / 20 = 2500 links per second. In this case, reducing the link pitch will increase the link cutting speed. However, the net throughput gained from a minor pitch reduction of a few percent are insignificant. Device designs should not be compromised in an effort to gain a small reduction in pitch.

LSD uses an off-line pre-processor called "sortlinks" to analyze the placement of links on a device design. A file is generated that contains the X, Y coordinates of every possible link on the device. The pre-processor reads the file then arranges the links into groups based on a very detailed model of the M3XX system dynamics. The output of the pre-processor is a file that assigns a group # and position in the group to every possible X, Y link location. The pre-processor also defines the group starting location, direction, pitch and number of links. This information is used during device processing to map repair information into Linksprint-compatible group # and position within the group format. The mapping from customer-supplied XY coordinate information into Linksprint compatible format is performed automatically by LSD supplied software, the user does need to understand the mapping process.

The pre-processor is able to find groupings that may appear to be counter-intuitive, but upon closer examination, prove to be time optimal. For instance, if sortlinks finds 2 links separated by less than about 200 microns, it will form a group of 2 links from the pair. While they may not appear to the designer to be related, sortlinks "knows" that they can be processed faster as a group of 2 links than as 2 individual links. Individual links that cannot be grouped with any others are processed as isolated "stop and cut" links (stop and cut is faster for 1 link, Linksprint is faster for 2 or more links separated by less than 200 microns). When the actual device is processed, the Linksprint algorithm examines the actual repair data given for particular group. The "sprint" will be terminated after the last cut link in a group without waiting to sweep across any remaining links that don't require cutting. If a group contains only 1 link to cut, then the link will be processed using a "stop and cut" move to save time.

The important consideration for a device designer interested in maximizing laser throughput is that 2 groups of links that share a common pitch and are spaced an even multiple of link pitch apart will be processed faster than if the groups are

spaced "off pitch". The sortlinks pre-processor will group the 2 banks into 1 longer bank all on a common pitch. The "blank space" between the groups will be filled in by sortlinks with phantom links.

Spot Size Considerations

There is a persistent desire amongst device designers to operate the laser at the smallest possible spot size. The prevailing logic assumes that the smaller spot is necessary to achieve finer link pitch and thereby shrink the device size. The smaller device allows more devices to be produced on a wafer with an overall reduction in the device cost to manufacture. Finer design rules with each generation of lithography equipment have enabled smaller devices to be produced - thus it is natural to expect that smaller spot sizes are the path to finer laser link pitches.

The logic fails to account for a few important facts:

The **goal** should be stated in economic terms - maximize the number of **good** devises per wafer;

Laser wavelength remains **constant** due to laser / material interaction physics. The benefit of shorter wavelength to improve resolution (ie: spot size) that enables the optical lithography process, does not apply with the laser process.

The relationship between spot size and total device yield is a complex function of laser fusing yield, total number of links to cut, number of die to repair and total die per wafer. The optimum link pitch, spot size and other operating parameters should be determined after careful consideration of all the controlling factors.

If the goal is stated in economic terms, then the focus of attention will be on achieving the best cost of manufacture for a device. If the goal is stated in terms of minimizing spot size, then sub-optimal economic operating point might be achieved in the process of optimizing the spot size. The problem lies in the "simple" assumption that minimum spot size is the best when actually minimum **cost** is best.

The reduction in wavelength with each successive generation of optical lithography equipment is critical to maintaining a suitable processing window despite the finer resolution. Laser fusing systems require a constant 1.047 micron wavelength to achieve high reliability of the fusing process. Attempts to use a frequency doubled YAG laser (0.524 micron wavelength) in the mid-1980's in an effort to achieve better resolution (i.e. smaller spot sizes) without sacrificing depth of focus were unsuccessful. The problem with shorter wavelength is increased absorption of the laser pulse into the silicon substrate as

the wavelength is reduced. This can cause substrate cracking with an attendant reliability issue.

With laser wavelength fixed at 1.047 microns, finer resolution must come at the expense of depth of focus. This is independent of laser machine design - it is related solely to the physics of light. Depth of focus varies as the **square** of the spot size, halving the spot size reduces the depth focus to one-quarter of the original value. Laser systems can be made to operate with very small depth of focus but the systems can be very difficult to operate and maintain in production environments. Most users have found the increased system reliability, uptime and device yield due to a larger spot size are more economically attractive than struggling with a smaller spot size.

If the goal is minimum device cost and the smallest spot size is not necessarily the best, then what spot size is best? The answer can only be found by a combination of analysis and experimentation. The process is more complex than can be fully discussed in this paper, however the basic approach is as follows. A spreadsheet is recommended to model the interaction of the many factors. The model looks at the number of raw die per wafer, the number of die to repair, the number of links per die and the number of links expected to be cut. The influence of spot size is modeled by a probability of successful fusing that is a function of spot size and device design rules. The fusing probability can only be reliably determined through extensive testing using suitably fabricated test structures. The probability of fusing a single link is raised to the power equal to the number of links to be cut:

$$\text{Die\_yield} = \text{link-yield}^{\#\_\text{links\_cut}}$$

As the number of links to be cut increases, the die yield suffers rapidly with any decrease in link_yield. On the other hand, if only a few links are expected to be cut, but many links are placed on the die due to layout considerations, then a small reduction in fusing yield due to smaller link pitch may be economically justified. The smaller link pitch affects all die, not just the repaired die and the smaller pitch is used in many places on each die. Most DRAM and SRAM designs exhibit a very small percentage of die area that is sensitive to laser processing design rules. A much larger fraction of die area is related to redundancy, but this area is not sensitive to laser design rules. The small laser-related percentage implies that a large change in link geometry is required to make a discernible impact on total device area. Even small changes in device geometry, however, may adversely impact fusing yield. The decrease in fixed to attempt (FTA) of repaired devices caused by use of tighter design rules may not be offset by the increased number of die that can be placed on a wafer.

The model can only produce useful design guidelines if the relationship between design rules and fusing yield is credible. Practice has shown that customer-specific link structures can have a significant impact on the fusing yield for a

given design rule. Proximity of non-fused structures, oxide thickness and material, the use on anti-reflective coatings, etc. make the use of "industry" guidelines of limited value. This implies that best results can be obtained by experimentally determining the fusing yield as a function of design rules. LSD can assist with the experimental process and developing a spreadsheet model of actual devices. Test structures must be provided by the customer as well as a means of assessing fusing yield (i.e. tester programs for the test structures). Industry practice, to the extent that it exists, is to operate with spot size (measured at $1/e^2$ - see below) of from 5.5 to 8 microns.

In summary, spot size isn't the issue, device yield and hence cost should be the driving factors in establishing design rules and system operating parameters. Careful analysis of the relationship between the many factors that relate the laser process to device cost will determine the optimum processing conditions.

One final comment about spot size needs to be made. Spot size is typically, but not universally, specified at the $1/e^2$ point. This point is chosen based on conventions typically used amongst optical engineers when describing the performance of lenses and related optical devices. **It does not mean the size of the hole made by a laser blast.** The actual "kerf" size may be larger or smaller than the theoretical spot size based on many device-specific factors. One other convention for specifying spot size uses the full width, half-maximum point (FWHM). The FWHM specification is related to the $1/e^2$ specification as follows:

spot_size(FWHM) = 0.58 ($1/e^2$) is equivalent to 2.3 microns when specified at FWHM.

Die Alignment Targets

The location and structures of alignment targets can influence device yield and, in some cases throughput. Die alignment is performed by laser scanning selected sites at reduced laser energy and monitoring the reflected laser intensity with a photo-detector inside the beam positioner. The reflected intensity varies as the laser spot is swept across a structure with significant variations in reflectivity. Dedicated aluminum targets are frequently used for die alignment. Targets placed at the die periphery provide locations from which X, Y offset, rotation (Theta) and X, Y scale factors can be determined.

The M3XX relies on die alignment following every movement of the XY table. The XY table used in the M3XX exhibits stepping errors of about 10 microns and the wafer itself can have small errors between die when measured across the entire wafer. The M3XX die alignment corrects for these error sources by aligning the working laser beam directly to the actual die just before the die is processed.

The M3XX has the capability to align once across an area spanning multiple die then process all the die in the area at once. The advantage of Multi-Die Align (MDA) becomes significant in applications where many die can be processed with a single alignment. This is particularly true with RF ID Tags and other small devices that are programmed, trimmed or repaired with a laser. Larger devices such as DRAMs and SRAMs benefit to a lesser extent from MDA. MDA cannot be used with devices that require on-line probing with the laser. The probe card pins obscure the adjacent die. We refer to the area processed by MDA as a "reticle". The analogy with a stepper reticle that exposes multiple die with a single exposure is obvious. Note: the laser can be programmed to follow the same reticle pattern as the stepper. Typically, the MDA reticle bears no relationship to the stepper reticle.

Alignment target design rules are the same regardless of the use of MDA. However, the placement of the targets at each die can be influenced by the use of MDA. First we will examine the structure of individual targets, then look at their placement a the die periphery.

The laser scanning process is non-destructive. The laser energy is set well below the cutting threshold during the scan. The minimum feature that can be located is a single "black/white" transition: where "black" is a non-reflective area and "white" is a reflective area such as an aluminum feature. Usually the substrate is considered non-reflective for the purpose of laser scanning. The minimum feature consists of a suitably large area of aluminum placed in an otherwise bare area of substrate. Information about location of the target can only be determined from the **transition** (edge) between the aluminum and the substrate. Small imperfections in the edge of the aluminum along with random "noise" in the optical, mechanical and electrical components of the laser system add a level of uncertainty to the location of any edge scan. The uncertainty is typically very small (under 0.1μ), but overall system accuracy can be improved by reducing the edge scan uncertainty.

Uncertainty can be reduced by spatial and temporal averaging. Spatial averaging is achieved at a single target by selecting targets with multiple edges. Each edge in the target scan provides additional information which can be used to reduce uncertainty. Scanning over an aluminum line, for example, provides 2 edges - a "LO-HI" transition followed by a "HI-LO" transition. Scanning over 2 suitably spaced aluminum lines provides 4 edges for a further reduction in edge location uncertainty. Multiple edges provide **spatial** averaging. Multiple scans over the target structure, separated by a user-set distance, provide both temporal and spatial averaging. Minor local imperfections in the target structure cause less of an error when multiple scans across multiple edges are averaged by the system software. Random (i.e. uncorrected) noise that might corrupt a single scan is reduced by averaging with other scans. The M3XX will allow up to

8 edges in a single target scan; 2 or 4 edges are the most commonly used targets.

The linewidths used for laser scanning targets should be a minimum of 2 times the focused spot size. Very narrow targets will not reflect all the available spot energy which will reduce the contrast in the scan. Linewidths wider than about 2 times the spot size do not add any value - all the information comes from the **edges** of the target, not the "HI" regions or the 'LO" regions.

The target material should be aluminum. Avoid the use of polymide over the mark. The region surrounding the mark should be free of any reflective materials.

Typical alignment marks are shown in Figure 1. Note that the targets for X and Y scans do not have to be located at the same point. Three sizes of targets are shown in the figures. The largest target, "A", is scanned first. At the time of the first scan the uncertainty in the die location is greatest and the tolerances surrounding the target are necessarily largest. Information from the first X and Y scans is used by the M3XX software to correct the location of subsequent scans. The subsequent targets can be smaller due to the reduction in uncertainty of die location. The targets shown are **minimum** sizes. If real-estate considerations allow for slightly larger targets then larger targets are preferred. Most users make all targets the same size as a convenience.

Target location can affect laser accuracy. Optimum accuracy is achieved when the targets are placed at the periphery of the area containing the links. Usually, the targets are placed at the four corners of the die thereby surrounding the entire die area. If all the links are located in a small region of the die then there might be some merit in placing the targets around the sub-region of the die with the links. The situation to avoid is placing the targets around a sub-region of the die and then **extrapolating** to some other region of the die to process links. If all the links lie within the target region then the error caused by a bad alignment mark will be no larger than the error of the mark itself. If the alignment is extrapolated to outside the region the error due to the bad target can be magnified by the extent of the extrapolation. Fortunately, the most common place to put the marks is at the corners of a die or even in the streets between die.

The number of targets to use depends on the desired level of accuracy, the use of MDA and rea- estate considerations. Typical memory repair applications use 3 or 4 pairs of targets placed in 3 or 4 corners of the die (1pair is 1 X and 1 Y scan). Four pairs placed 1 pair to a die corner is the preferred arrangement.

Very small die that are processed using MDA require only a single pair of targets located anywhere in or around the die. The M3XX MDA software will use target pairs from 4 die located at the periphery of the MDA reticle for die alignment.

Larger die such as those commonly found on SRAM and DRAM devices should include a minimum of 3 pairs, preferably 4 pairs, of targets even if MDA is expected to be used.

Reflectivity of the aluminum is important for reliable alignment operations. The oxide over the alignment targets should be less than 10,000 Å for best results.

Wafer Alignment

Wafers must be aligned to the M3XX coordinate system before any processing can begin. Wafer placement errors are caused by robotic or manual loading inaccuracy as well as any pattern shift relative to the wafer periphery. Two methods are available in the M3XX to align wafers: a laser scanning alignment method and an (optional) pattern recognition method. The laser scanning method is included in the system for historical purposes. The pattern recognition method is the preferred method due to increased speed, reliability and ease of setup. Only the pattern recognition method will be discussed here.

The M3XX includes a CCTV system in the beam positioner that looks through the same optical path as the laser. The viewing path is steered by the same beam steering components that position the laser spot over the working field. The pattern recognition option adds a vision processor to the M3XX electronics that receives the image from the CCTV camera, performs image processing functions on the image and re-transmits the image to a monitor located adjacent to the system workstation monitor. The CCTV monitor provided a real-time image of the die at all times.

Wafer alignment searches for a unique pattern that best matches a previously stored template within a user-set area on the die. The application engineer trains the pattern recognition system on a template during the initial device setup process. The goal is to find an area on the die, usually 150 to 500 microns square, that exhibits a very unique patterns of lines. Since most wafers exhibit predominately vertical and horizontal structures, a template with circular, diagonal or otherwise unique features is desirable. The company logo is often used since most logos exhibit the desired unique characteristics.

Unique patterns located on test die are used on occasion. The pattern must be located at multiple test die to allow the pattern recognition system to resolve wafer offset and theta. Theta is resolved by template matching at 2 or more locations separated by an integer number of die. If the template is found on all die then the choice of how many die to use in the baseline is not significant. If the template is only found in the test die area, then the exact number of die between test sites must be used for the baseline dimension.

The M3XX uses a monochromatic source of illumination for the CCTV optical path. The single wavelength will reflect off the many layers that are found in a typical wafer causing constructive or destructive interference in the image. The presence of the interference simply makes some structures dark, others bright in the image. The contrast is desirable, in that it provides features that the pattern recognition system can identify. The interference effects can become a problem when the layer thickness cannot be controlled very well and the pattern of dark/bright changes from wafer to wafer. If the stored template includes areas that change significantly from wafer to wafer or batch to batch, then the pattern recognition system will have difficulty matching the stored template to the actual wafers. The application engineer can create solutions to this situation but the solution typically requires more time to perform the wafer alignment operation.

The preferred solution is to use a template area that has "strong" aluminum features. The aluminum provides a strong reflection relative to most other layers and will therefore provide dominant bright features in the image despite changes in reflectivity from other layers. A company logo with bold features is ideal. A very "busy logo" with many fine details is less desirable due to the reduced variation in reflectivity compared to the lower metal layers (s). The fewer layers that cover the target metal and the less variation in the covering layers, the greater the reliability and speed of the pattern recognition process.

The problem with large contrast variations is usually only significant in fabs that are using older processing equipment with less stringent control of layer thickness. Most current generation DRAM and SRAM fabs produce parts with sufficient layer control so that the problem of varying contrast is eliminated or at least not significant.

Wafer Handling

The M3XX product line incorporates an optional cassette to cassette robotic wafer loader - the H955. The H955 is used in virtually all applications where wafer processing time is less than 30 minutes per wafer or where cleanliness is important. The H955 is designed for operation in a Class 10 cleanroom and is compatible with most SEMI-compatible cassette automation schemes. The H955 reduces the chance of operator error (such as selecting the wrong wafer from a cassette) and reduces manual handling which may contaminate wafers.

The H955 employs an integrated robotics mechanism that fetches, pre-aligns and loads wafers onto the M3XX wafer stage. The pre-aligner uses an optical sensor that is scanned beneath the periphery of the wafer to determine the wafer center and orientation features. One or more flats or a single notch are used to establish water orientation. Wafer sizes from 100 to 200 mm are accommodated.

SEMI-specified cassettes are compatible with the locating method used on the H955. The "H-bar" feature is used for cassette registration per SEMI specifications. Most non-standard cassettes can also be accommodated with a slight modification to the locating plates. The modification, if required, can be done at the factory or in the field in most instances.

The H955 robot uses dual paddles to reduce wafer exchange time with the M3XX. After the initial wafer is loaded all other wafers are fetched and pre-aligned while the previous wafer is being processed. The dual paddle allows the robot to unload the finished wafer and load the next wafer in a single motion sequence. The processing of the new wafer begins while the finished wafer is returned to the cassette. The only impact on cycletime is the actual wafer exchange time - approximately 17 seconds.

The H955 uses sophisticated vacuum sensing on each wafer paddle, the pre-align station and the M3XX chucktop. Wafers are detected based on the vacuum sensing. All wafer transfers require the appropriate vacuum detection or an error is reported. The robot also uses vacuum sensing to automatically determine the proper height for all transfer operations. For instance, the first time the robot fetches a wafer from a cassette, the robot approaches the theoretical pickup height slowly while monitoring the paddle vacuum. The robot records the height at which the vacuum is first detected. The measured height is used as an offset when fetching all subsequent wafers from the cassette.

The M3XX application software can be configured to handle cassettes with missing wafers. Normally the H955 will only load wafers from cassette slots that are explicitly specified in the repair data file for the lot. If a wafer is missing the data for that wafer will not be present in the repair data file. The software can also be set-up to search for wafers in a cassette. In this mode, the robot searches each slot in succession. If a wafer is detected, the software then searches for the corresponding repair data file (based on the slot number). If the data file is not located the wafer is marked in the status file and the wafer is not processed.

Optical Character Recognition

Optical Character Recognition (OCR) can be used to provide an additional level of security to insure that the proper repair information is used to process a wafer. The OCR option consists of an optical unit incorporating a camera that mounts to the H955 handler and character reading software that is downloaded into the same vision processor that is also used for wafer alignment. The H955 handler software automatically presents the wafer to the optical unit after the pre-alignment operation. The vision processor analyzes the image and returns a character string to the application program corresponding to the observed wafer ID. The wafer ID string incorporates a checksum character that is used by the OCR software to identify errors in the reading process. It is very unlikely that a

read error will pass the checksum test and be returned to the applications as a "false positive"

The preferred OCR mark conforms to SEMI font specifications. Wafer ID marks must be placed on the top side and near to the wafer edge. The mark can be located at any orientation relative to the wafer flat (or notch) and the mark can be oriented either towards or away from the wafer center. Marks created by a laser marking process ("hard" marks) are read with the greatest reliability. Wafer ID marks created by patterning within the wafer layers ("soft" marks) are more difficult to read reliably. The patterning process produces a lower contrast mark and the contrast will vary with changes in layer thickness. The OCR optical unit will work with soft marks but the setup process is more difficult than with hard marks.

Multiple modes of operation are available to the application program to handle the wafer ID string. Typically the repair data file will include the ID strings for each wafer. The application program will verify that the proper ID is read from the wafer located in a given slot. If the wafer ID does not match the repair file for the particular slot , no attempt is made to match the wafer ID with any other repair file. The assumption is that if any inconsistency is found between the repair information and the actual cassette then processing should be stopped until the inconsistency is eliminated through operator intervention.

The alternate mode of operation will allow processing to continue if any repair file can be found that matches the observed wafer ID. This mode allows processing to continue in the event a wafer is misplaced in a cassette. In some situations this may be considered acceptable. This mode of operation can also be used to allow continuous wafer processing and automatic changeover form 1 device type to another. The repair file can include the device type as a data field. The application program will read the device type from the repair file and make the appropriate changes to setup files when the type of device changes.

When placed in this mode of operation, the system can run without any operator intervention beyond simply placing and removing cassettes from the handler. The H955 will detect the presence of a cassette, fetch a wafer, read the wafer ID and search for the device type and repair information. After each wafer is completed the repair file is marked to indicate the wafer has been processed.

When all wafers in cassette 1 are processed the H955 will start working from cassette 2 (if present). The operator may replace cassette 1 with a new cassette of wafers at any time while the robot is using cassette 2. The robot will automatically begin processing the new cassette when processing of cassette 2 is completed.

Repair Data File

Repair information is passed to the M3XX laser system as a data file. The problem of passing repair information from a tester to the laser has two components: the physical connection and the file format.

The preferred means of passing a file containing repair information from a test to the laser is via the Ethernet network interface built-in to the laser workstation. The system software makes the file transfer process transparent to the operator. The network interface eliminates the problem of handling disks or other media in the cleanroom.

Many facilities already incorporate an Ethernet link between various processing equipment. Adding the laser to the network can usually be done by the local system administrator with a minimum of difficulty.

Some facilities mange the tester data using a commercial data management software package such as Teradyne's FIRMS. LSD has interfaced to FIRMS and other data management systems. The interface is usually straight forward and can be easily implemented by LSD's application engineers. Most commercial data packages provide a set of calling routines that can be used by the application program to access the tester data.

Other means of passing repair information from the tester to the laser have been implemented including RS232 and IEEE communications, 3.5" floppy disks and ¼" cartridges tape. If floppy disks are used either DOS or UNIX TAR format are recommended. In some cases a dedicated PC may be used to read a non-standard format or media and then communicate the file information to the laser workstation via Ethernet or RS232.

The format of the data file varies with each customer location. Usually, the tester or other processing equipment determines the organization of the data files. LSD has created a process that easily accommodates almost any file format. LSD typically creates a simple translation program that reads the customer-specific file and converts the data into a format that the laser application can use. The translation process also converts the data to a form that is compatible with the notion of groups required by Linksprint (see above).

A typical file format is an ASCII file that lists the wafer ID, which die are to be repaired and the link coordinates of the links to cut. In some cases the customer prefers to have the repair algorithm reside with the laser rather than the tester. In this case the repair file will simply indicate the row or column address to use for repair of a particular die. The translation program that LSD creates will convert the row/column address into a link cut list based on a customer-provided repair algorithm.

The amount of data required to specify the repair of a DRAM or SRAM wafer is typically small compared to the file capabilities of a modern workstation. The use

of ASCII files, while less than optimal from a computation standpoint, is clearly preferable when debugging problems. The repair file can be easily reviewed, printed or modified with a text editor to test special cases or resolve a problem. The computation required by the laser workstation to read the repair file is insignificant for all but the most special of cases.

If the laser being used to program a die then the number of links to cut on a wafer can become very large. In a typical ID Tag programming example, a 4" wafer may hold 5000 die and each die may require 100 links to cut. Since each die is unique the file for a single wafer will contain about 500,000 coordinate pairs of data. In such cases an alternative encoding scheme for the "repair" data is often preferred. The choice of format is usually application-specific. If the data is sequential then the laser can be provided with a starting value and an algorithm for sequencing. In this case the file is reduced to simply the starting value. In other cases a binary file may be used with bitmap representing the links to cut. The laser application program cam be easily modified by LSD to handle these situations.

Applications Program

LSD provides a turnkey application with virtually every system delivered. The "generic" application program has evolved over many years of installing laser processing solutions to incorporate a large set of features and capabilities. Minor customization of the program is typically required to handle customer-specific issues: file format, user interface conventions, special reports, etc. LSD typically customizes the application program to handle the special cases as part of a normal system installation.

The generic application program is intended to be used by the customer setup person(s) to setup, debug and support production processing of multiple device types. The program makes extensive use of the graphical user interface capabilities of the workstation to simplify the setup person's task. The setup process for a new device involves defining device-specific information such as die size, wafer align templates, alignment mark location and link coordinate data. All device-specific data is automatically stored in device data files on the system workstation.

The generic application program can be used by operators to control production operations of the laser. The application program monitors the login file of the operator and defeats certain options from the application menu that are inappropriate for production operators. The login file for the setup person enables the full functionality of the program.

LSD will typically create a custom production menu at each location to provide a very simple operations menu for production functions. In most cases the menu

can be configured to require the absolute minimum of keystrokes and display only the data required by the operator to monitor production operations. Access to the full functionality of the setup software is controlled by the login file. Production operators are limited to accessing only the topmost level of functionality; the setup person is given full access to the software capabilities.

Throughput

Throughput is entirely dependent upon device characteristics, processing time per die may vary by a factor of 10 or more due to die characteristics. For instance, a logic programming application may require 15 seconds or more per die to cut over 5000 links. An RF ID Transponder may require less than 0.2 seconds to process when the link layout is optimized for laser throughput. LSD has developed a detailed computer model of system performance for the purposes of estimating processing time. The model considers device characteristics such as wafer size, die size, percentage to be repaired, average number of links cut per die, pitch and layout of links groups, MDA, etc. The model has proven to be accurate over a broad range of applications.

The model is intended for use by LSD engineering personnel to estimate throughput of hypothetical device layouts. The model is also used by LSD application engineers when analyzing an application for possible throughput improvements. The model clearly shows the percentage in throughput improvement that will result from various changes to the application. Knowing in advance the relative improvement from various activities helps focus engineering efforts where the most benefit will be realized.

LSD will gladly analyze customer-submitted device information for the purpose of estimating throughput.